



## INEXACT NEWTON METHOD FOR M-TENSOR EQUATIONS

DONG-HUI LI, HONG-BO GUAN\* AND JIE-FENG XU

**Abstract:** We first investigate properties of M-tensor equations. In particular, we show that if the constant term of the equation is nonnegative, then finding a nonnegative solution of the equation can be done by finding a positive solution of a lower dimensional M-tensor equation. We then propose an inexact Newton method to find a positive solution to the lower dimensional equation and establish its global convergence. We also show that the convergence rate of the method is quadratic. At last, we do numerical experiments to test the proposed Newton method. The results show that the proposed Newton method has a very good numerical performance.

**Key words:** *M-tensor equation, inexact Newton method, global convergence, quadratic convergence*

**Mathematics Subject Classification:** *65H10, 65K10, 90C30*

### 1 Introduction

Tensor equation is a special system of nonlinear equations. It is also called multilinear equation. Tensor equation can be expressed as

$$F(x) = \mathcal{A}x^{m-1} - b = 0, \quad (1.1)$$

where  $x, b \in \mathbb{R}^n$  and  $\mathcal{A}$  is an  $m$ th-order  $n$ -dimensional tensor that takes the form

$$\mathcal{A} = (a_{i_1 i_2 \dots i_m}), \quad a_{i_1 i_2 \dots i_m} \in \mathbb{R}, \quad 1 \leq i_1, i_2, \dots, i_m \leq n,$$

and  $\mathcal{A}x^{m-1} \in \mathbb{R}^n$  with elements

$$(\mathcal{A}x^{m-1})_i = \sum_{i_2, \dots, i_m} a_{i i_2 \dots i_m} x_{i_2} \dots x_{i_m}, \quad i = 1, 2, \dots, n.$$

For convenience of presentation, we introduce some concepts and notations, which will be used throughout the paper. We denote the set of all  $m$ th-order  $n$ -dimensional tensors by  $\mathcal{T}(m, n)$ . We first introduce the concepts of Z-matrix and M-matrix.

---

\*Corresponding author. Supported by the NSF of China grant No.11771157, 11801184, Guangdong Basic and Applied Basic Research Foundation grant No.2020B1515310013 and Hunan Provincial Education Department of China grant No.20C0559

**Definition 1.1** ([3]). A matrix  $A$  is called a Z-matrix if all its off-diagonal entries are nonpositive. It is apparent that a Z-matrix  $A$  can be written as

$$A = sI - B,$$

where  $B$  is a nonnegative matrix ( $B \geq 0$ ) and  $s > 0$ ; When  $s \geq \rho(B)$ , we call  $A$  an M-matrix; And further when  $s > \rho(B)$ , we call  $A$  a nonsingular M-matrix, where  $\rho(B)$  is the spectral radius of  $B$ .

The concept of M-tensor is an extension of the definition of M-matrix. Now we introduce the definition of M-tensor and other structure tensors that will be involved in this paper.

**Definition 1.2.** [6, 7, 8, 19, 25, 26, 27, 36] Let  $\mathcal{A} \in \mathcal{T}(m, n)$ .

- $\mathcal{A}$  is called a nonnegative tensor, denoted by  $\mathcal{A} \geq 0$ , if all its elements are nonnegative, i.e.,  $a_{i_1 i_2 \dots i_m} \geq 0$ ,  $\forall i_1, \dots, i_m \in [n]$ , where  $[n] = \{1, 2, \dots, n\}$ .
- $\mathcal{A}$  is called a symmetric tensor, if its elements  $a_{i_1 i_2 \dots i_m}$  are invariant under any permutation of their indices. In particular, for every index  $i \in [n]$ , if an  $(m-1)$ th-order  $n$ -dimensional square tensor  $\mathcal{A}_i := (a_{i i_2 \dots i_m})_{1 \leq i_2, \dots, i_m \leq n}$  is symmetric, then  $\mathcal{A}$  is called semi-symmetric tensor with respect to the indices  $\{i_2, \dots, i_m\}$ . The set of all  $m$ th-order  $n$ -dimensional symmetric tensors is denoted by  $\mathcal{ST}(m, n)$ .
- $\mathcal{A}$  is called the identity tensor, denoted by  $\mathcal{I}$ , if its diagonal elements are all ones and other elements are zeros, i.e., all  $a_{i_1 i_2 \dots i_m} = 0$  except  $a_{i i \dots i} = 1$ ,  $\forall i \in [n]$ .
- If a real number  $\lambda$  and a nonzero real vector  $x \in \mathbb{R}^n$  satisfy

$$\mathcal{A}x^{m-1} = \lambda x^{[m-1]},$$

then  $\lambda$  is called an H-eigenvalue of  $\mathcal{A}$  and  $x$  is called an H-eigenvector of  $\mathcal{A}$  associated with  $\lambda$ . Here  $x^{[m-1]} = (x_1^{m-1}, x_2^{m-1}, \dots, x_n^{m-1})^T$ .

- $\mathcal{A}$  is called an M-tensor, if it can be written as

$$\mathcal{A} = s\mathcal{I} - \mathcal{B}, \quad \mathcal{B} \geq 0, \quad s \geq \rho(\mathcal{B}), \quad (1.2)$$

where  $\rho(\mathcal{B})$  is the spectral radius of tensor  $\mathcal{B}$ , that is

$$\rho(\mathcal{B}) = \max \{|\lambda| : \lambda \text{ is an eigenvalue of } \mathcal{B}\}.$$

If  $s > \rho(\mathcal{B})$ , then  $\mathcal{A}$  is called a strong or nonsingular M-tensor.

- $\mathcal{A}$  is called a lower triangular tensor, if its possibly nonzero elements are  $a_{i_1 i_2 \dots i_m}$  with  $i_1 = 1, 2, \dots, n$  and  $i_2, \dots, i_m \leq i_1$  and all other elements of  $\mathcal{A}$  are zeros.  $\mathcal{A}$  is called a strictly lower triangular tensor, if its possibly nonzero elements are  $a_{i_1 i_2 \dots i_m}$  with  $i_1 = 1, 2, \dots, n$  and  $i_2, \dots, i_m < i_1$  and all other elements of  $\mathcal{A}$  are zeros.
- $\mathcal{A}$  is called reducible if there is an index set  $I \subset [n]$  such that the elements of  $\mathcal{A}$  satisfy

$$a_{i_1 i_2 \dots i_m} = 0, \quad \forall i_1 \in I, \forall i_2, \dots, i_m \notin I.$$

If  $\mathcal{A}$  is not reducible, then we call  $\mathcal{A}$  irreducible.

The notation  $\mathcal{A}x^m$  will denote the homogenous polynomial of degree  $m$ , i.e.,

$$\mathcal{A}x^m = x^T \mathcal{A}x^{m-1} = \sum_{i_1, \dots, i_m} a_{i_1 \dots i_m} x_{i_1} \dots x_{i_m}.$$

In the case  $\mathcal{A} \in \mathcal{ST}(m, n)$ , the derivative of the homogeneous polynomial  $\mathcal{A}x^m$  can be expressed as  $\nabla(\mathcal{A}x^m) = m\mathcal{A}x^{m-1}$ .

In the case  $\mathcal{A}$  is semi-symmetric, it holds that

$$\nabla(\mathcal{A}x^{m-1}) = (m - 1)\mathcal{A}x^{m-2},$$

where  $\mathcal{A}x^{m-2}$  is an  $n \times n$  matrix defined by

$$(\mathcal{A}x^{m-2})_{ij} = \sum_{i_3, \dots, i_m} a_{ij i_3 \dots i_m} x_{i_3} \dots x_{i_m}, \quad i, j = 1, 2, \dots, n.$$

In the definition of reducible tensor, the index set  $I \subset [n]$  can be arbitrary. In our paper, we will need some special reducible tensor where the index set  $I$  is contained in some specified set. For the sake of convenience, we make a slight extension to the definition of reducible tensors.

**Definition 1.3.** Tensor  $\mathcal{A} \in \mathcal{T}(m, n)$  is called reducible with respect to  $I \subset [n]$  if its elements satisfy

$$a_{i_1 i_2 \dots i_m} = 0, \quad \forall i_1 \in I, \forall i_2, \dots, i_m \notin I.$$

It is easy to see that tensor  $\mathcal{A} \in \mathcal{T}(m, n)$  is reducible if and only if there is an index set  $I \subset [n]$  such that it is reducible respect to  $I$ .

We call index pair  $(I, I_c)$  a partition of  $[n]$  if  $I, I_c \subset [n]$ ,  $I \cap I_c = \emptyset$  and  $I \cup I_c = [n]$ .

For  $x, y \in \mathbb{R}^n$  and  $\alpha \in \mathbb{R}$ , the notations  $x \circ y$  and  $x^{[\alpha]}$  are vectors in  $\mathbb{R}^n$  defined by

$$x \circ y = (x_1 y_1, \dots, x_n y_n)^T$$

and

$$x^{[\alpha]} = (x_1^\alpha, \dots, x_n^\alpha)^T$$

respectively.

For an  $x \in \mathbb{R}^n$ , and an index set  $I \subseteq [n]$ , we denote by  $x_I$  the subvector of  $x$ , whose elements are  $x_i, i \in I$ . For matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  and index sets  $I, J \subseteq [n]$ , we denote by  $A_{IJ}$  the submatrix of  $A$  with elements  $a_{ij}, i \in I, j \in J$ . The principle submatrix  $A_{II}$  is abbreviated as  $A_I$ . For tensor  $\mathcal{A} = (a_{i_1 \dots i_m}) \in \mathcal{T}(m, n)$ , we denote by  $\mathcal{A}_I$  the principle subtensor of  $\mathcal{A}$  with elements  $a_{i_1 \dots i_m}, i_1, \dots, i_m \in I$ .

We use  $\mathbb{R}_+^n$  and  $\mathbb{R}_{++}^n$  to denote the sets of all nonnegative vectors and positive vectors in  $\mathbb{R}^n$ . That is,

$$\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\} \quad \text{and} \quad \mathbb{R}_{++}^n = \{x \in \mathbb{R}^n \mid x > 0\}.$$

If  $\mathcal{A}$  is an M-tensor, we call the tensor equation an M-tensor equation and abbreviate it as M-Teq.

The following theorem comes from [3, 8, 11].

**Theorem 1.4.** Let  $\mathcal{A} \in \mathcal{T}(m, n)$ .

- ([8]) If  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_{++}^n$ , then the M-Teq (1.1) has a unique positive solution.

- ([11]) *If  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_+^n$ , then the M-Teq (1.1) has a nonnegative solution.*
- ([3]) *For a Z-matrix  $A \in \mathbb{R}^{n \times n}$ , the following statements are equivalent.*
  - (i)  *$A$  is a nonsingular M-matrix.*
  - (ii)  *$Av \in \mathbb{R}_{++}^n$  for some vector  $v \in \mathbb{R}_{++}^n$ .*
  - (iii) *All the principal minors of  $A$  are positive.*

Tensor equation is also called multilinear equation. It appears in many practical fields including data mining and numerical partial differential equations [5, 8, 9, 10, 14, 15, 16, 32]. The study in numerical methods for solving tensor equations has begun only a few years ago. Most existing methods focus on solving the M-Teq under the restriction  $b \in \mathbb{R}_{++}^n$  or  $b \in \mathbb{R}_+^n$ . Such as the iterative methods in [8], the homotopy method in [12], the tensor splitting method in [20], the Newton-type method in [13], the continuous time neural network method in [28], the preconditioned tensor splitting method in [22], the preconditioned SOR method in [21], the preconditioned Jacobi type method in [35], the nonnegativity preserving algorithm in [2]. There are also a few methods that can solve M-Teq (1.1) without restriction  $b \in \mathbb{R}_{++}^n$  or that  $\mathcal{A}$  is an M tensor. Those methods include the splitting method by Li, Guan and Wang [18], and Li, Xie and Xu [15], the alternating projection method by Li, Dai and Gao [17], the alternating iterative methods by Liang, Zheng and Zhao [23] etc.. Related works can also be found in [4, 5, 16, 24, 29, 30, 31, 32, 33, 34].

Newton's method is a well-known efficient method for solving nonlinear equations. An attractive property of the method is its quadratic convergence rate. However, in many cases, the standard Newton method may fail to work or lose its quadratic convergence property when applied to solve tensor equation (1.1). We refer to [18] for details.

Recently, He, Ling, Qi and Zhou [13] proposed a Newton type method for solving the M-Teq (1.1) with  $b \in \mathbb{R}_{++}^n$ . Unlike the standard Newton method for solving nonlinear equations, by utilizing the special structure of the equation (1.1), the authors transformed the equation into an equivalent form through a variable transformation  $y = x^{[m]}$ . Starting from some positive initial point, the method generates a sequence of positive iterates. An attractive property of the method is that the Jacobian matrices of the equation at the iterates are nonsingular. As a result, the method is well defined and retains the global and quadratic convergence. The reported numerical results in [13] confirmed the quadratic convergence property of that method.

It should be pointed out that the positivity of  $b$  plays an important role in the Newton method by He, Ling, Qi and Zhou [13]. It is not known if the method in [13] is still well defined and preserves quadratic convergence property if there is some  $i$  satisfying  $b_i = 0$ . The purpose of this paper is to develop a Newton method to find a nonnegative solution of the equation (1.1) with  $b \in \mathbb{R}_+^n$ . Our idea is similar to but different from that of the method in [13]. Specifically, we will reformulate the equation via the variable transformation  $y = x^{[m-1]}$ . Such an idea comes from the following observation. Consider a very special tensor equation

$$Ax^{[m-1]} - b = 0,$$

corresponding to the tensor equation (1.1) where the only nonzero elements of  $\mathcal{A}$  are  $a_{ij\dots j} = a_{ij}$ ,  $i, j = 1, 2, \dots, n$ . For that special equation, the tensor equation is equivalent to the system of linear equation  $Ay - b = 0$  with  $y = x^{[m-1]}$ . As a result, the corresponding Newton method terminates at a solution of the equation within one iteration. Another difference

between our method and the method in [13] is that we will consider the equation (1.1) with  $b \in \mathbb{R}_+^n$ . The case where  $b$  has zero elements causes the problem be much more difficult. Existing techniques that deal with equation (1.1) with  $b \in \mathbb{R}_{++}^n$  may fail. To overcome that difficulty, we will propose a criterion that can identify the zero elements in a nonnegative solution of the M-tensor equation. From a computational view point, the criterion is easy to implement. By the use of that criterion, we can get a nonnegative solution of the M-tensor equation (1.1) by finding a positive solution to a lower dimensional M-tensor equation with nonnegative constant term.

Based on that criterion, we propose a Newton method for finding a positive solution of the M-Teq with  $b \in \mathbb{R}_+^n$  and establish its global and quadratic convergence.

The remainder of the paper is organized as follows. In the next section, we investigate some nice properties of the M-tensor equation (1.1). In particular, we propose a criterion to distinguish zero and nonzero elements of a nonnegative solution of the equation. In Section 3, we propose a Newton method to get a positive solution to the M-Teq (1.1) with  $b \in \mathbb{R}_{++}^n$  and establish its global and quadratic convergence. In Section 4, we extend the method proposed in Section 3 to the M-Teq (1.1) with  $b \in \mathbb{R}_+^n$  and show its global and quadratic convergence. At last, we do numerical experiments to test the proposed method in Section 5.

## 2 Properties of M-Tensor Equations

Throughout this section, we suppose that tensor  $\mathcal{A} \in \mathcal{T}(m, n)$  is a strong M-tensor.

The following lemma was proved by Li, Guan and Wang [18].

**Lemma 2.1.** *If  $\mathcal{A}$  is a strong M-tensor, and the feasible set  $\mathcal{S}$  defined by*

$$\mathcal{S} \triangleq \{x \in \mathbb{R}_+^n \mid F(x) = \mathcal{A}x^{m-1} - b \leq 0\}$$

*is not empty, then  $\mathcal{S}$  has a largest element that is the largest nonnegative solution to the M-tensor equation  $F(x) = \mathcal{A}x^{m-1} - b = 0$ .*

As an application of the last lemma, we have the following proposition.

**Proposition 2.2.** *Let  $\mathcal{A}$  be a strong M-tensor and  $b^{(1)}, b^{(2)} \in \mathbb{R}^n$  satisfy  $b^{(2)} \geq b^{(1)}$ . Suppose that the M-tensor equation*

$$\mathcal{A}x^{m-1} - b^{(1)} = 0 \tag{2.1}$$

*has a nonnegative solution  $x^{(1)}$ . Then the M-tensor equation*

$$\mathcal{A}x^{m-1} - b^{(2)} = 0 \tag{2.2}$$

*has a nonnegative solution  $x^{(2)}$  satisfying  $x^{(2)} \geq x^{(1)}$ . In particular, if  $b^{(1)} > 0$ , then the unique positive solution  $\bar{x}^{(1)}$  of (2.1) and the unique positive solution  $\bar{x}^{(2)}$  of (2.2) satisfies  $\bar{x}^{(2)} \geq \bar{x}^{(1)}$ .*

*Proof.* Define

$$\mathcal{S}_1 \triangleq \{x \in \mathbb{R}_+^n \mid \mathcal{A}x^{m-1} - b^{(1)} \leq 0\}$$

and

$$\mathcal{S}_2 \triangleq \{x \in \mathbb{R}_+^n \mid \mathcal{A}x^{m-1} - b^{(2)} \leq 0\}.$$

Since  $b^{(1)} \leq b^{(2)}$ , we obviously have

$$\mathcal{S}_1 \subseteq \mathcal{S}_2.$$

By the assumption that (2.1) has a nonnegative solution. By Lemma 2.1 the set  $\mathcal{S}_1$  is nonempty and has a largest element  $\bar{x}^{(1)}$  that is a solution to the equation (2.1). Consequently, the set  $\mathcal{S}_2$  is nonempty and has a largest element  $x^{(2)}$  that is a solution to the equation (2.2). It is clear that

$$x^{(2)} \geq \bar{x}^{(1)} \geq x^{(1)}.$$

If  $b^{(1)} > 0$ , then the unique positive solution  $\bar{x}^{(1)}$  is the largest element of  $\mathcal{S}_1$  and  $\bar{x}^{(2)}$  is the largest element of  $\mathcal{S}_2$ . As a result, we have  $\bar{x}^{(2)} \geq \bar{x}^{(1)}$ . The proof is completed.  $\square$

**Theorem 2.3.** *Suppose that  $\mathcal{A}$  is a strong M-tensor. Then the following statements are true.*

(i) *The tensor equation*

$$\mathcal{A}x^{m-1} = 0 \tag{2.3}$$

*has a unique solution  $x = 0$ .*

(ii) *If  $-b \in \mathbb{R}_+^n \setminus \{0\}$ , then the M-Teq (1.1) has no nonnegative solutions.*

(iii) *The following relationship holds*

$$x \circ \mathcal{A}x^{m-1} = 0 \iff x = 0.$$

(iv) *It holds that*

$$\lim_{\|x\| \rightarrow \infty} \|\mathcal{A}x^{m-1}\| = +\infty.$$

(v) *For any  $b \in \mathbb{R}^n$ , the solution set of the M-tensor equation (1.1), if not empty, is compact.*

*Proof.* Conclusion (i) is trivial because zero is not an eigenvalue of any strong M-tensor.

(ii) Suppose for some  $b \leq 0$ ,  $b \neq 0$ , the M-Teq (1.1) has a nonnegative solution  $\bar{x} \neq 0$ . Clearly,  $\bar{x} \geq 0$ . Denote  $I = \{i : \bar{x}_i > 0\}$ . Let  $\mathcal{D}$  be a diagonal tensor whose diagonal elements are  $d_{i\dots i} = -b_i \bar{x}_i^{-(m-1)}$ ,  $\forall i \in I$ , and  $d_{i\dots i} = 0$ ,  $\forall i \notin I$ . Let  $\bar{\mathcal{A}} = \mathcal{A} + \mathcal{D}$ . It is obvious that  $\bar{\mathcal{A}}$  is a strong M-tensor. Clearly,  $\bar{\mathcal{A}}_I$  is a strong M-tensor too. However, it holds that  $\bar{\mathcal{A}}_I \bar{x}_I^{m-1} = 0$ , which yields a contradiction.

Conclusion (iii) follows from (i) directly because any principal subtensor of a strong M-tensor is a strong M-tensor.

(iv) Suppose on the contrary that there is some sequence  $\{x_k\}$  satisfying  $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$  such that the sequence  $\{\|\mathcal{A}x_k^{m-1}\|\}$  is bounded. Then we have

$$\lim_{k \rightarrow \infty} \frac{\|\mathcal{A}x_k^{m-1}\|}{\|x_k\|^{m-1}} = 0.$$

Let  $y_k = x_k/\|x_k\|$  and  $\bar{y}$  be an accumulation point of the sequence  $\{y_k\}$ . It is easy to see that  $\bar{y} \neq 0$  but  $\mathcal{A}\bar{y}^{m-1} = 0$ , which contradicts with (i).

The conclusion (v) is a direct corollary of the conclusion (iv).  $\square$

The latter part of this section focuses on the M-Teq (1.1) with  $b \in \mathbb{R}_+^n$ . We denote

$$I^+(b) = \{i : b_i > 0\}, \quad \text{and} \quad I^0(b) = \{i : b_i = 0\}.$$

We first show the following theorem.

**Theorem 2.4.** *Suppose that  $\mathcal{A}$  is irreducible and is a strong M-tensor. Then every non-negative solution of the M-Teq (1.1) with  $b \in \mathbb{R}_+^n$  must be positive.*

*Proof.* Suppose on the contrary that the M-Teq (1.1) with  $b \in \mathbb{R}_+^n$  has a nonnegative solution  $\bar{x}$  satisfying  $I = \{i \mid \bar{x}_i = 0\} \neq \emptyset$ . We have for any  $i \in I$ ,

$$0 = \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \bar{x}_{i_2} \dots \bar{x}_{i_m} - b_i = \sum_{\{i_2, \dots, i_m\} \subseteq I_c} a_{ii_2 \dots i_m} \bar{x}_{i_2} \dots \bar{x}_{i_m} - b_i \leq 0.$$

Since  $\bar{x}_j > 0, \forall j \in I_c$ , the last inequality yields  $b_i = 0$  and

$$a_{ii_2 \dots i_m} = 0, \quad \forall i_2, \dots, i_m \notin I.$$

It shows that tensor  $\mathcal{A}$  is reducible with respect to  $I$ , which yields a contradiction. □

By the proof of the last theorem, we have the following corollary.

**Corollary 2.5.** *Suppose that  $\mathcal{A}$  is a strong M-tensor. If the M-Teq (1.1) with  $b \in \mathbb{R}_+^n$  has a nonnegative solution  $\bar{x}$  with zero elements, then  $\mathcal{A}$  is reducible with respect to some  $I \subseteq I^0(b)$ .*

The following theorem characterizes a nonnegative solution of the M-Teq (1.1) with  $b \in \mathbb{R}_+^n$ .

**Theorem 2.6.** *Suppose that  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_+^n$ . Then the M-Teq (1.1) has a nonnegative solution with zero elements if and only if  $\mathcal{A}$  is reducible with respect to some  $I \subseteq I^0(b)$ . Moreover, for a nonnegative solution  $\bar{x}$  of the M-Teq (1.1),  $\bar{x}_i = 0$  iff  $i \in I$ .*

*Proof.* The “only if” part follows from Corollary 2.5 directly.

Suppose that tensor  $\mathcal{A}$  is reducible with respect to some  $I^1 \subseteq I^0(b)$ . It is easy to see that the M-tensor equation (1.1) has a solution  $\bar{x}$  with  $\bar{x}_{I^1} = 0$ . Denote  $I_c^1 = [n] \setminus I^1$ . Consider the lower dimension M-tensor equation

$$\mathcal{A}_{I_c^1} x_{I_c^1}^{m-1} - b_{I_c^1} = 0. \tag{2.4}$$

Since  $b_{I_c^1} \geq 0$ , the last equation has a nonnegative solution  $\bar{x}_{I_c^1}$ , which together with  $\bar{x}_{I^1} = 0$  forms a nonnegative solution to the M-tensor equation (1.1).

If  $\mathcal{A}_{I_c^1}$  is irreducible, then  $I = I^1$  is the desired index set. Otherwise,  $\mathcal{A}_{I_c^1}$  is reducible with respect to some  $I^2 \subset I_c^1$  satisfying  $I^2 \subseteq I^0(b_{I_c^1})$ . We consider the lower dimensional M-tensor equation (2.4). Following a similar discussion to the above process, we can get another lower dimensional tensor equation whose nonnegative solution together with some zeros forms a nonnegative solution to (1.1). Continuing this process finitely many times, we can get a desirable index set  $I \subset I^0(b)$ . □

**Remark 2.7.** The above theorem provides a way to reduce the size of an M-tensor equation with  $b \in \mathbb{R}_+^n$ . Specifically, in the case tensor  $\mathcal{A}$  is reducible with respect to some  $I \subseteq I^0(b)$ , we can get a solution to (1.1) by finding a positive solution to the lower dimensional tensor equation

$$\mathcal{A}_{I_c} x_{I_c}^{m-1} - b_{I_c} = 0,$$

where  $I_c = [n] \setminus I$ .

As a direct corollary of the last theorem, we have the following results, which gives a necessary and sufficient condition for the M-tensor equation (1.1) with  $b \in \mathbb{R}_+^n$  to have a positive solution.

**Corollary 2.8.** *Suppose that  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_+^n$ . Then there is an index set  $I \subseteq I^0(b)$  such that every nonnegative solution to the following lower dimensional tensor equation with  $I_c = [n] \setminus I$*

$$\mathcal{A}_{I_c} x_{I_c}^{m-1} - b_{I_c} = 0$$

*is positive. Moreover, the positive solution  $x_{I_c}$  of the last equation together with  $x_I = 0$  forms a nonnegative solution to the M-Teq (1.1).*

The following lemma gives another interesting property for an M-Teq.

**Lemma 2.9.** *Suppose that  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_+^n$ . Suppose further that every nonnegative solution of the M-Teq (1.1) is positive. Then there is an index set  $J \subseteq I^0(b)$  such that for each  $i \in J$ , there is at least one  $i_j \notin J$ ,  $2 \leq j \leq m$  such that  $a_{ii_2 \dots i_m} \neq 0$ .*

*Proof.* Let  $J_0 = I^0(b)$ . It is easy to see that there must be at least one  $i \in J_0$  and at least one  $i_j \notin J_0$ ,  $2 \leq j \leq m$  such that  $a_{ii_2 \dots i_m} \neq 0$ . Otherwise, the M-Teq has a nonnegative solution  $\tilde{x}$  with  $\tilde{x}_i = 0$ ,  $\forall i \in J_0$ , with yields a contradiction.

If  $J_0$  does not meet the requirement, we get an index set  $J_1 \subset J_0$  consisting of all indices  $i \in J_0$  that does not meet the requirement. If  $J_1$  still is not the desired index set, we can further get a small index set  $J_2 \subset J_1$ . We repeat the process. At last, we get the desired index set  $J$ .  $\square$

Based on the last lemma, we can show the nonsingularity property of the Jacobian  $F'$  at the positive solutions.

**Theorem 2.10.** *Suppose that the M-Teq (1.1) with a strong M-tensor  $\mathcal{A}$  and  $b \in \mathbb{R}_+^n$  has a positive solution  $\bar{x}$ . Then the Jacobian  $F'(\bar{x})$  is a nonsingular M-matrix. Let  $f(y) = F(y^{[\frac{1}{m-1}]})$  and  $\bar{y} = \bar{x}^{[m-1]}$ . Then  $f'(\bar{y})$  is also a nonsingular M-matrix.*

*Proof.* It is easy to derive for any  $y > 0$ ,

$$f'(y) = \mathcal{A}(y^{[\frac{1}{m-1}]})^{m-2} \text{diag}(y^{[\frac{1}{m-1}-1]}).$$

It shows that the nonsingularity of  $f'(\bar{y})$  is the same as the nonsingularity of  $F'(\bar{x})$ .

Let  $J \subseteq I^0(b)$  be the index set specified by Lemma 2.9 and  $I = [n] \setminus J$ . Write the Jacobian matrix  $F'(\bar{x})$  as the block form

$$F'(\bar{x}) = (m-1)\mathcal{A}\bar{x}^{m-2} = \begin{pmatrix} A_{II} & A_{IJ} \\ A_{JI} & A_{JJ} \end{pmatrix}.$$

Since  $\bar{x}$  is a positive solution of (1.1), it follows from Lemma 2.9 that  $A_{JI}$  has no zero rows.

That  $\bar{x}$  is a solution to (1.1) yields  $F'(\bar{x})\bar{x} = (m-1)b$ . Writing it as block form, we get

$$\begin{cases} A_{II}\bar{x}_I + A_{IJ}\bar{x}_J = (m-1)b_I, \\ A_{JI}\bar{x}_I + A_{JJ}\bar{x}_J = (m-1)b_J. \end{cases} \quad (2.5)$$

It follows from the last equality of the above system that

$$A_{JJ}\bar{x}_J = (m-1)b_J - A_{JI}\bar{x}_I = -A_{JI}\bar{x}_I > 0.$$

Since  $A_{JJ}$  is a Z-matrix, the last inequality implies that  $A_{JJ}$  is a nonsingular M-matrix. By Theorem 1.4, it is easy to see that the Schur complement  $A_{II} - A_{IJ}A_{JJ}^{-1}A_{JI}$  is a nonsingular M-matrix.

If  $J = I^0(b)$ , then  $I = I^+(b)$ . We get from the first equality of (2.5),

$$(A_{II} - A_{IJ}A_{JJ}^{-1}A_{JI}) \bar{x}_I = (m - 1)b_I > 0.$$

Clearly, matrix  $A_{II} - A_{IJ}A_{JJ}^{-1}A_{JI}$  is a Z-matrix. Consequently, the last inequality shows that the Schur complement of  $A_{JJ}$  is a nonsingular M-matrix too. Therefore,  $A = F'(\bar{x})$  is a nonsingular M-matrix.

In the case  $J \subset I^0(b)$ , we denote  $J_1 = J$ ,  $I_1 = I$  and  $A_1 = A_{I_1I_1} - A_{I_1J_1}A_{J_1J_1}^{-1}A_{J_1I_1}$ . Then to show  $F'(\bar{x})$  is a nonsingular M-matrix is equivalent to show that the lower dimensional Z-matrix  $A_1$  is a nonsingular M-matrix. It is clear that  $\bar{x}_{I_1}$  satisfies the lower dimensional system of linear equations

$$A_1\bar{x}_{I_1} = (m - 1)b_{I_1}.$$

Similar to above arguments, we can get a partition  $(I_2, J_2)$  to the index set  $I_1$  that possesses the same properties as  $(I_1, J_1)$ . Repeat the process finitely many times, we can get  $J_t = I^0(b_{I_{t-1}})$ . As a result, we can verify that  $F'(\bar{x})$  is a nonsingular M-matrix.  $\square$

### 3 A Newton Method for M-Tensor Equation (1.1) with $b \in \mathbb{R}_{++}^n$

In this section, we propose a Newton method to find the unique positive solution to (1.1) with  $b \in \mathbb{R}_{++}^n$ . Throughout this section, without specification, we always suppose that the following assumption holds.

**Assumption 1** Tensor  $\mathcal{A}$  is a semi-symmetric and strong M-tensor, and  $b \in \mathbb{R}_{++}^n$ .

Recently, He, Ling, Qi and Zhou [13] developed a Newton method for solving the M-Teq (1.1) with  $b \in \mathbb{R}_{++}^n$ . By making a variable transformation  $x = y^{[\frac{1}{m}]}$ , they formulated the equation as the following equivalent nonlinear equation:

$$W(y) = D(y) \cdot F(y^{[\frac{1}{m}]}) = D(y) \cdot \mathcal{A}(y^{[\frac{1}{m}]})^{m-1} - D(y) \cdot b = 0,$$

where  $D(y) = \text{diag}(y_i^{\frac{1}{m}-1})$ . The above equation has some nice properties such as the nonsingularity of the Jacobian  $W'(y)$  for any  $y > 0$ . In the case where  $\mathcal{A}$  is symmetric, the tensor equation (1.1) is the stationary equation of the minimization problem

$$\min \bar{f}(y) = \frac{1}{m} \mathcal{A}(y^{[\frac{1}{m}]})^m - b^T(y^{[\frac{1}{m}]})$$

because the gradient of  $\bar{f}(y)$  is

$$\nabla \bar{f}(y) = \frac{1}{m} W(y) = \frac{1}{m} D(y) \cdot \nabla f(y^{[\frac{1}{m}]})$$

In what follows, we propose a Newton method for finding the unique positive solution of the M-Teq (1.1). Our idea to develop the Newton method is similar to but different from that in [13]. Details are given below.

Since our purpose is to get a positive solution of the M-Teq (1.1), we restrict  $x \in \mathbb{R}_{++}^n$ . Making a variable transformation  $y = x^{[m-1]}$ , we formulate the M-Teq (1.1) as

$$f(y) = F(y^{[\frac{1}{m-1}]}) = \mathcal{A}(y^{[\frac{1}{m-1}]})^{m-1} - b = 0. \tag{3.1}$$

In the case  $m = 2$ , the last system remains the system of linear equations unchanged.

A direct computation gives

$$f'(y) = \mathcal{A}\left(y^{\lfloor \frac{1}{m-1} \rfloor}\right)^{m-2} \text{diag}\left(y^{\lfloor \frac{1}{m-1} \rfloor - 1}\right).$$

It follows that

$$f'(y)y = \mathcal{A}\left(y^{\lfloor \frac{1}{m-1} \rfloor}\right)^{m-2} \text{diag}\left(y^{\lfloor \frac{1}{m-1} \rfloor - 1}\right)y = \mathcal{A}\left(y^{\lfloor \frac{1}{m-1} \rfloor}\right)^{m-1} = f(y) + b.$$

Clearly, the positive solutions of the M-Teq (1.1) are positive solutions of the following nonlinear equation:

$$E(y) \triangleq \text{diag}\left(y^{[-1]}\right)f(y) = \left(y^{[-1]} \circ f(y)\right) = 0. \quad (3.2)$$

The Jacobian of  $E(y)$  is

$$\begin{aligned} E'(y) &= \text{diag}\left(y^{[-1]}\right)f'(y) - \text{diag}\left(f(y)\right)\text{diag}\left(y^{[-2]}\right) \\ &= \text{diag}\left(y^{[-1]}\right)\left[f'(y) - \text{diag}\left(f(y)\right)\text{diag}\left(y^{[-1]}\right)\right]. \end{aligned}$$

It is a non-symmetric Z-matrix. For any  $y > 0$ , it holds that

$$\begin{aligned} E'(y)y &= \text{diag}\left(y^{[-1]}\right)\left[f'(y)y - \text{diag}\left(f(y)\right)\left(y^{[-1]}\right)y\right] \\ &= \text{diag}\left(y^{[-1]}\right)\left[\mathcal{A}\left(y^{\lfloor \frac{1}{m-1} \rfloor}\right)^{m-1} - f(y)\right] \\ &= \text{diag}\left(y^{[-1]}\right)b > 0. \end{aligned}$$

Consequently, we have got the following proposition.

**Proposition 3.1.** *Let  $E : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$  be defined by (3.2). For any  $y > 0$ , the Jacobian  $E'(y)$  is a nonsingular M-matrix. Moreover, the equation (3.2) has a unique positive solution that is the unique positive solution to the M-Teq (1.1).*

We are going to develop a Newton method for solving the nonlinear equation (3.2) in which the Newton direction  $d_k$  is the solution to the system of linear equations

$$E'(y_k)d + E(y_k) = 0,$$

i.e.,

$$\text{diag}\left(y_k^{[-1]}\right)\left[f'(y_k) - \text{diag}\left(\frac{f(y_k)}{y_k}\right)\right]d + \text{diag}\left(y_k^{[-1]}\right)f(y_k) = 0,$$

or equivalently

$$\left[f'(y_k) - \text{diag}\left(\frac{f(y_k)}{y_k}\right)\right]d + f(y_k) = 0 \quad (3.3)$$

Here  $\text{diag}\left(\frac{f(y_k)}{y_k}\right)$  is a diagonal matrix whose diagonals are  $\frac{f_i(y_k)}{(y_k)_i}$ ,  $i = 1, 2, \dots, n$ . We can regard  $d_k$  as an inexact Newton direction for solving the equation  $f(y) = 0$  because the Newton equation (3.3) can be written as

$$f'(y_k)d_k + f(y_k) = r_k, \quad r_k = \text{diag}\left(\frac{f(y_k)}{y_k}\right)d_k = O(\|f(y_k)\| \|d_k\|),$$

if  $y_k > 0$  is bounded away from zero.

Let  $y_k(\alpha) = y_k + \alpha d_k$ . Then  $y_k(\alpha)$  satisfies

$$\left[ f'(y_k) - \text{diag} \left( \frac{f(y_k)}{y_k} \right) \right] y_k(\alpha) = \left[ f'(y_k) - \text{diag} \left( \frac{f(y_k)}{y_k} \right) \right] y_k - \alpha f(y_k) = b - \alpha f(y_k).$$

Since the Jacobian

$$E'(y_k) = \text{diag} \left( y_k^{[-1]} \right) \left[ f'(y_k) - \text{diag} \left( \frac{f(y_k)}{y_k} \right) \right]$$

is an M-matrix and  $y_k > 0$ , it is clear that the matrix

$$f'(y_k) - \text{diag} \left( \frac{f(y_k)}{y_k} \right)$$

is an M-matrix too. Therefore, the inequality  $y_k(\alpha) > 0$  will be guaranteed if

$$b - \alpha f(y_k) > 0. \tag{3.4}$$

Let

$$\bar{\alpha}_k^{\max} = \min \left\{ \frac{b_i}{f_i(y_k)} : f_i(y_k) > 0 \right\}. \tag{3.5}$$

It is clear that

$$y_k + \alpha d_k > 0, \quad \forall \alpha \in (0, \bar{\alpha}_k^{\max}).$$

The iterative process of the Newton method is stated as follows.

**Algorithm 1 Inexact Newton’s Method**

**Initial.** Given constants  $\sigma, \rho \in (0, 1)$  and  $\epsilon > 0$ . Select an initial point  $x_0 > 0$ . such that  $y_0 = x_0^{[m-1]}$  satisfies  $f(y_0) < b$ . Let  $k = 0$ .

**Step 1.** Stop if  $\|E(y_k)\| < \epsilon$ .

**Step 2.** Solve the system of linear equations (3.3) to get  $d_k$ .

**Step 3.** For given constant  $\sigma \in (0, 1)$ , let  $\alpha_k = \max\{\rho^i : i = 0, 1, \dots\}$  such that  $y_k + \alpha_k d_k > 0$  and that the inequality

$$\|E(y_k + \alpha_k d_k)\|^2 \leq (1 - 2\sigma\alpha_k)\|E(y_k)\|^2, \quad \sigma \in (0, 1). \tag{3.6}$$

is satisfied.

**Step 3.** Let  $y_{k+1} = y_k + \alpha_k d_k$ . Go to Step 1.

**Remark 3.2.** (1) The difference between the above method and the standard Newton method lies in Step 3 where an addition restriction exists. Specifically, to generate positive iterates, we need to restrict the steplength  $\alpha_k > 0$  sufficiently small such that  $x_{k+1}$  is positive for any  $k \geq 0$ . Such restriction make the analysis in the global convergence and the quadratic convergence will be different from that for the standard Newton method.

(2) It is easy to see that the inequality (3.4) is guaranteed if  $f(y_k) < b$ . So, at the beginning, we select  $y_0 > 0$  satisfying  $f(y_0) < b$  and at each iteration, we let  $y_{k+1} = y_k + \alpha_k d_k$  such that  $f(y_{k+1}) < b$ . In this way, the inequalities  $f(y_k) < b$  holds for all  $k$ .

**Lemma 3.3.** *Let  $\{y_k\}$  be generated by Algorithm 1. Then there is a positive constant  $c$  such that*

$$y_k \geq c\mathbf{e}, \quad \forall k \geq 0, \tag{3.7}$$

where  $\mathbf{e} = (1, 1, \dots, 1)^T$ .

*Proof.* It is clear that the sequence of the function evaluations  $\{\|E(y_k)\|\}$  is decreasing and hence bounded by some constant  $\overline{M} > 0$ , i.e.,

$$\|E(y_k)\| \leq \overline{M}.$$

Since  $\mathcal{A}$  is an M-tensor, there is a constant  $s > 0$  and a nonnegative tensor  $\mathcal{B} \geq 0$  such that  $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ , where  $\mathcal{I}$  stands for the identity tensor whose diagonals are all ones and all other elements are zeros.

By the definition of  $E(y)$ , we have

$$E(y) = s\mathbf{e} - \text{diag}(y^{-1})\mathcal{B}\left(y^{\lfloor \frac{1}{m-1} \rfloor}\right)^{m-1} - b \circ (y^{\lfloor -1 \rfloor}).$$

Since  $\mathcal{B} \geq 0$ , the last inequality implies for any  $y > 0$  and each  $i \in [n]$

$$|E_i(y)| \geq \frac{b_i}{y_i} + y_i^{-1} \left( \mathcal{B}\left(y^{\lfloor \frac{1}{m-1} \rfloor}\right)^{m-1} \right)_i - s \geq \frac{b_i}{y_i} - s.$$

Suppose there is an index  $i$  and an infinite set  $K$  such that  $\lim_{k \in K, k \rightarrow \infty} (y_k)_i = 0$ . We have

$$\overline{M} \geq \lim_{k \in K, k \rightarrow \infty} |E_i(y_k)| \geq \lim_{k \in K, k \rightarrow \infty} \frac{b_i}{(y_k)_i} - s = +\infty,$$

which yields a contradiction. The contradiction shows that the inequality in (3.7) is satisfied with some positive constant  $c$ . □

The following theorem establishes the global convergence of the proposed method.

**Theorem 3.4.** *Suppose that the sequence  $\{y_k\}$  generated by Algorithm 1 is bounded. Then  $\{y_k\}$  converges to the unique positive solution to the M-Teq (1.1).*

*Proof.* We first show that the maximum step length  $\bar{\alpha}_k^{\max}$  satisfying (3.5) can be bounded away from zero. That is, there is a constant  $\bar{\alpha}$  such that

$$\bar{\alpha}_k^{\max} \geq \bar{\alpha}, \quad \forall k \geq 0. \tag{3.8}$$

Indeed, it follows from the last lemma that

$$\overline{M} \geq |E_i(y_k)| = \frac{|f_i(y_k)|}{(y_k)_i}.$$

Since  $\{y_k\}$  is bounded, the last inequality implies that for each  $i$ ,  $\{|f_i(y_k)|\}$  is bounded too. By the definition of  $\bar{\alpha}_k^{\max}$ , it is bounded away from some constant  $\bar{\alpha}$ . Consequently, the inequality (3.8) is satisfied for all  $k \geq 0$ .

Next, we show that there is an accumulation point  $\bar{y}$  of  $\{y_k\}$  that is a positive solution to (1.1).

Suppose  $\{y_k\}_K \rightarrow \bar{y}$ . By Lemma 3.3, it is clear that  $\bar{y} > 0$ . Consequently,  $E'(\bar{y})$  is an M-matrix. Moreover,

$$\lim_{k \in K, k \rightarrow \infty} d_k = -E'(\bar{y})^{-1}E(\bar{y}) \triangleq \bar{d}.$$

Without loss of generality, we let  $\lim_{k \in K, k \rightarrow \infty} \alpha_k = \tilde{\alpha}$ .

If  $\tilde{\alpha} > 0$ , then the inequality (3.6) shows that  $\{\|E(y_{k+1})\|\}_k \rightarrow 0$ .

If  $\tilde{\alpha} = 0$ , then when  $k \in K$  is sufficiently large, the inequality (3.6) is not satisfied with  $\alpha'_k = \rho^{-1}\alpha_k$ , i.e.,

$$\|E(y_k + \alpha'_k d_k)\|^2 - \|E(y_k)\|^2 > -2\sigma\alpha'_k\|E(y_k)\|^2, \quad \sigma \in (0, 1).$$

Dividing both sides of the inequality by  $\alpha'_k$  and then taking limits as  $k \rightarrow \infty$  with  $k \in K$ , we get

$$-\|E(\bar{y})\|^2 = E(\bar{y})^T E'(\bar{y})\bar{d} \geq -2\sigma\|E(\bar{y})\|^2,$$

which implies  $E(\bar{y}) = 0$ .

Since  $\{\|E(y_k)\|\}$  converges, it follows from Lemma 3.3 that every accumulation point of  $\{y_k\}$  is a positive solution to (1.1). However, the positive solution of (1.1) is unique. Consequently, the whole sequence  $\{y_k\}$  converges to the unique positive solution to (1.1).  $\square$

By a standard argument, it is not difficult to show that the convergence rate of  $\{y_k\}$  is quadratic.

**Theorem 3.5.** *Let the conditions in Theorem 3.4 hold. Then the convergence rate of  $\{y_k\}$  is quadratic.*

#### 4 An Extension

In this section, we extend the Newton method proposed in the last section to the M-Teq (1.1) with  $b \in \mathbb{R}_+^n$ . In the case  $b$  has zero elements, the M-Teq may have multiple nonnegative or positive solutions. Our purpose is to find one nonnegative or positive solution of the equation.

We see from the definition of  $E(y)$  that the function  $E(y)$  and its Jacobian are not well defined at a point with zero elements. Therefore, the Newton method proposed in the last section can not be applied to find a nonnegative solution with zero elements. Fortunately, from Corollary 2.8, we can get a nonnegative solution of (1.1) by finding a positive solution to a lower dimensional M-Teq.

Without loss of generality, we make the following assumption.

**Assumption 2** Suppose that tensor  $\mathcal{A}$  is a semi-symmetric and strong M-tensor, and  $b \in \mathbb{R}_+^n$ . Moreover, every nonnegative solution of the M-Teq (1.1) is positive.

Similar to the Newton method by He, Ling, Qi and Zhou [13], we propose another Newton method, which we call a regularized Newton method, such that the method is still globally and quadratically convergent without assuming the boundedness of the generated sequence of iterates.

It is easy to see that the M-Teq (1.1) is equivalent to the following nonlinear equation

$$E(t, y) \triangleq \begin{pmatrix} t \\ y^{[-1]} \circ f(y) + ty \end{pmatrix} = \begin{pmatrix} t \\ \bar{E}(t, y) \end{pmatrix} = 0, \tag{4.1}$$

where

$$\bar{E}(t, y) = E(y) + ty = y^{[-1]} \circ f(y) + ty.$$

The Jacobian of  $E(t, y)$  is

$$E'(t, y) = \begin{pmatrix} 1 & 0 \\ y & \bar{E}'_y(t, y) \end{pmatrix},$$

where

$$\bar{E}'_y(t, y) = E'(y) + tI$$

satisfying

$$\bar{E}'_y(t, y)y = E'(y)y + ty = y^{[-1]} \circ b + ty > 0, \quad \forall y \in \mathbb{R}_{++}^n, \quad \forall t > 0.$$

Since  $\bar{E}'_y(t, y)$  is a Z-matrix, the last inequality shows that it is a nonsingular M-matrix. As a result, for any  $t > 0$  and any  $y \in \mathbb{R}_{++}^n$ , the Jacobian  $E'(t, y)$  is nonsingular.

Now, we propose a Newton method for solving the equivalent nonlinear equation (4.1) to the M-Teq (1.1). The idea is similar to the Newton method by He, Ling, Qi and Zhou [13]. Details are given below.

Given constant  $\gamma \in (0, 1)$ . Denote

$$\theta(t, y) = \frac{1}{2} \|E(t, y)\|^2, \quad \beta(t, y) = \gamma \min\{1, \|E(t_k, y_k)\|^2\}.$$

The subproblem of the method is the following system of linear equations:

$$E'(t_k, y_k)d_k + E(t_k, y_k) = \beta(t_k, y_k)\mathbf{e}_1, \quad (4.2)$$

where  $\mathbf{e}_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^{n+1}$ . Let  $d_k = (d_k^t, d_k^y)$ .

Suppose  $t_k \leq \bar{t}$  with  $\bar{t}$  satisfying  $\bar{t}\gamma < 1$ . Then the Newton direction  $d_k$  satisfies

$$\begin{aligned} \nabla\theta(t_k, y_k)^T d_k &= E(t_k, y_k)^T E'(t_k, y_k)d_k = -\|E(t_k, y_k)\|^2 + \beta(t_k, y_k) \\ &\leq -(1 - \gamma\bar{t})\|E(t_k, y_k)\|^2. \end{aligned} \quad (4.3)$$

As a result, for given constant  $\sigma \in (0, 1)$ , the following inequality

$$\theta(t_k + \alpha_k d_k^t, y_k + \alpha_k d_k^y) \leq [1 - 2\sigma(1 - \gamma\bar{t})\alpha_k]\theta(t_k, y_k) \quad (4.4)$$

is satisfied for all  $\alpha_k > 0$  sufficiently small.

The steps of the method are stated as follows.

### Algorithm 2 Regularized Newton Method

**Initial.** Given constants  $\gamma, \sigma, \rho \in (0, 1)$ ,  $\epsilon > 0$  and  $\bar{t} > 0$  such that  $\bar{t}\gamma < 1$ . Given initial point  $x_0 > 0$  and  $t_0 = \bar{t}$ . Let  $y_0 = x_0^{[m-1]}$  and  $k = 0$ .

**Step 1.** Stop if  $\|E(t_k, y_k)\| \leq \epsilon$ .

**Step 2.** Solve the system of linear equations (4.2) to get  $d_k$ .

**Step 3.** Find  $\alpha_k = \max\{\rho^i : i = 0, 1, \dots\}$  such that  $y_k + \rho^i d_k^y > 0$  and that (4.4) is satisfied with  $\alpha_k = \rho^i$ .

**Step 4.** Let  $y_{k+1} = y_k + \alpha_k d_k^y$  and  $t_{k+1} = t_k + \alpha_k d_k^t$ .

**Step 5.** Let  $k := k + 1$ . Go to Step 1.

Following a similar argument as the proof of Lemma 3.2 of [13], it is not difficult to get the following proposition. It particularly shows that the above algorithm is well-defined.

**Proposition 4.1.** *Suppose that  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_+^n$ . Then the sequence of iterates  $\{(t_k, y_k)\}$  generated by Algorithm 2 satisfies*

$$0 < t_{k+1} \leq t_k \leq \bar{t}$$

and

$$t_k > \bar{t}\beta(t_k, y_k).$$

In addition, the sequence of function evaluations  $\{\theta(t_k, y_k)\}$  is decreasing.

Since  $\mathcal{A}$  is an M-tensor, there is a constant  $s > 0$  and a nonnegative tensor  $\mathcal{B} = (b_{i_1 \dots i_m})$  such that  $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ , where  $\mathcal{I}$  is the identity tensor whose diagonal entries are all ones and all other elements are zeros. By the definition of  $E(y)$ , it is easy to get

$$E(y) = \mathbf{se} - y^{[-1]} \circ \mathcal{B} \left( y^{\left[\frac{1}{m-1}\right]} \right)^{m-1} - y^{[-1]} \circ b,$$

where  $\mathbf{e} = (1, 1, \dots, 1)^T$ .

**Lemma 4.2.** *Suppose that  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_+^n$ . Then the sequence of iterates  $\{y_k\}$  generated by Algorithm 2 is bounded away from zero. In other words, there is a constant  $\eta > 0$  such that*

$$(y_k)_i \geq \eta, \quad \forall k \geq 0, \forall i = 1, 2, \dots, n.$$

*Proof.* Suppose that there is an index  $i$  and a subsequence  $\{y_k\}_K$  such that  $\lim_{k \rightarrow \infty, k \in K} (y_k)_i = 0$ . Without loss of generality, we suppose  $\{y_k\}_K \rightarrow \bar{y}$ , where some elements of  $\bar{y}$  may be  $+\infty$ . Denote  $I = \{i : \bar{y}_i = 0\}$  and  $I_c = [n] \setminus I_c$ . Since Let  $\{\theta(t_k, y_k)\}$  is decreasing, it is bounded and so is the sequence  $\{\|E(t_k, y_k)\|\}$ . Let  $C > 0$  be an upper bound of the sequence  $\{\|\bar{E}(t_k, y_k)\|\}$ .

For each  $i \in I$ , it holds that

$$\begin{aligned} C &\geq |\bar{E}_i(t_k, y_k)| = \left| \frac{1}{(y_k)_i} \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \left( (y_k)_{i_2}^{\frac{1}{m-1}} \dots (y_k)_{i_m}^{\frac{1}{m-1}} \right) - \frac{b_i}{(y_k)_i} + t_k (y_k)_i \right| \\ &= \left| s - \frac{1}{(y_k)_i} \sum_{i_2, \dots, i_m} b_{ii_2 \dots i_m} \left( (y_k)_{i_2}^{\frac{1}{m-1}} \dots (y_k)_{i_m}^{\frac{1}{m-1}} \right) - \frac{b_i}{(y_k)_i} + t_k (y_k)_i \right| \\ &\geq \sum_{i_2, \dots, i_m} b_{ii_2 \dots i_m} \left( \frac{(y_k)_{i_2}}{(y_k)_i} \dots \frac{(y_k)_{i_m}}{(y_k)_i} \right)^{\frac{1}{m-1}} + \frac{b_i}{(y_k)_i} - t_k (y_k)_i - s \\ &\geq \sum_{i_2, \dots, i_m \in I_c} b_{ii_2 \dots i_m} \left( \frac{(y_k)_{i_2}}{(y_k)_i} \dots \frac{(y_k)_{i_m}}{(y_k)_i} \right)^{\frac{1}{m-1}} + \frac{b_i}{(y_k)_i} - t_k (y_k)_i - s. \end{aligned}$$

Notice that for any  $i \in I_c$ ,  $\bar{y}_i > 0$ . Since  $t_k \leq \bar{t}$  and  $(y_k)_i \rightarrow 0$ , as  $k \rightarrow \infty$  with  $k \in K$ , the last inequality implies  $b_i = 0$  and  $a_{ii_2 \dots i_m} = b_{ii_2 \dots i_m} = 0, \forall i_2, \dots, i_m \in I_c$ . It means that tensor  $\mathcal{A}$  is reducible with respect to index set  $I$ . It then follows from Theorem 2.6 that the M-Teq (1.1) has a nonnegative solution that has zero elements. This is a contradiction which shows that  $\{y_k\}$  is bounded away from zero.  $\square$

**Lemma 4.3.** *Suppose that  $\mathcal{A}$  is a strong M-tensor and  $b \in \mathbb{R}_+^n$ . If there is a  $\tilde{t} > 0$  such that  $t_k \geq \tilde{t}$  for all  $k \geq 0$ , then the sequence of iterates  $\{y_k\}$  generated by Algorithm 2 is bounded.*

*Proof.* Denote by  $i_k$  the index satisfying  $(y_k)_{i_k} = \|y_k\|_\infty$ . Since  $\{\theta(t_k, y_k)\}$  has an upper bound, so is  $\{\|\bar{E}(t_k, y_k)\|\}$ . Let  $C$  be an upper bound of  $\{\|\bar{E}(t_k, y_k)\|\}$ . It is clear that

$$\left| \sum_{i_2, \dots, i_m} a_{i_k i_2 \dots i_m} \left( \frac{(y_k)_{i_2}}{(y_k)_{i_k}} \dots \frac{(y_k)_{i_m}}{(y_k)_{i_k}} \right)^{\frac{1}{m-1}} \right| \leq \sum_{i_2, \dots, i_m} |a_{i_k i_2 \dots i_m}| \triangleq \tilde{a}_{i_k}$$

is bounded. Therefore, we obtain

$$\begin{aligned} C &\geq \|\bar{E}(t_k, y_k)\| \\ &\geq \left| \frac{1}{(y_k)_{i_k}} \sum_{i_2, \dots, i_m} a_{i_k i_2 \dots i_m} \left( (y_k)_{i_2}^{\frac{1}{m-1}} \dots (y_k)_{i_m}^{\frac{1}{m-1}} \right) - \frac{b_{i_k}}{(y_k)_{i_k}} + t_k (y_k)_{i_k} \right| \\ &\geq t_k (y_k)_{i_k} - \tilde{a}_{i_k} - \frac{b_{i_k}}{(y_k)_{i_k}}. \end{aligned}$$

The last inequality together with  $t_k \geq \tilde{t}$  implies that  $\{\|y_k\|\}$  is bounded. □

The following theorem establishes the global convergence of Algorithm 2.

**Theorem 4.4.** *Suppose that  $\mathcal{A}$  is a strong  $M$ -tensor and  $b \in \mathbb{R}_+^n$ . Then every accumulation point of the sequence of iterates  $\{(t_k, y_k)\}$  generated by Algorithm 2 is a positive solution to the  $M$ -Teq (1.1).*

*Proof.* It suffices to show that the sequence  $\{\theta(t_k, y_k)\}$  converges to zero by contradiction. Suppose on the contrary that there is a constant  $\delta > 0$  such that  $\theta(t_k, y_k) \geq \delta, \forall k \geq 0$ . Then

$$\tilde{t} \triangleq \lim_{t \rightarrow \infty} t_k \geq \bar{t} \lim_{t \rightarrow \infty} \beta(t_k, y_k) \geq \bar{t} \gamma \min\{1, 2\delta\} > 0.$$

By Lemma 4.3,  $\{y_k\}$  is bounded. Let the subsequence  $\{y_k\}_K$  converge to some point  $\bar{y}$ . Lemma 4.2 ensures  $\bar{y} > 0$ . It is easy to show that the Jacobian  $E'(\bar{t}, \bar{y})$  is a nonsingular  $M$ -matrix. Consequently,  $\{d_k\}_K$  is bounded. Without loss of generality, we suppose  $\{d_k\}_K$  converges to some  $\bar{d}$ . Since  $\bar{y} > 0$ , there is a constant  $\alpha^{\min} > 0$  such that  $y_k + \alpha_k d_k > 0, \forall \alpha_k \in (0, \alpha^{\min})$ . Let  $\bar{\alpha} = \liminf_{k \rightarrow \infty, k \in K} \alpha_k$ . If  $\bar{\alpha} > 0$ , the line search condition (4.4) implies  $\theta(\bar{y}, \bar{t}) = 0$ . If  $\bar{\alpha} = 0$ , then when  $k$  is sufficiently large, the inequality (4.4) is not satisfied with  $\alpha'_k = \alpha_k \rho^{-1}$ , i.e.,

$$\theta(t_k + \alpha'_k d_k^t, y_k + \alpha'_k d_k^y) - \theta(t_k, y_k) \geq -2\sigma(1 - \gamma \bar{t}) \alpha'_k \theta(t_k, y_k).$$

Dividing both sides of the last inequality by  $\alpha'_k$  and then taking limits as  $k \rightarrow \infty$  with  $k \in K$ , we get

$$\nabla \theta(\bar{t}, \bar{y})^T \bar{d} \geq -2\sigma(1 - \gamma \bar{t}) \theta(\bar{t}, \bar{y}).$$

On the other hand, by taking limits in both sides of (4.3) as  $k \rightarrow \infty$  with  $k \in K$ , we obtain

$$\nabla \theta(\bar{t}, \bar{y})^T \bar{d} \leq -2(1 - \gamma \bar{t}) \theta(\bar{t}, \bar{y}).$$

Since  $\sigma \in (0, 1)$ , the last two inequalities implies  $\theta(\bar{y}, \bar{t}) = 0$ , which yields a contradiction. As a result, we conclude that  $\{\theta(t_k, y_k)\}$  converges to zero. The proof is complete. □

The last theorem has shown that every accumulation point is a positive solution to the  $M$ -Teq (1.1). However, it does not guarantee the existence of the accumulation point. The following theorem shows that the sequence  $\{y_k\}$  is bounded. As a result, it ensures the existence of the accumulation point.

**Theorem 4.5.** *Suppose that  $\mathcal{A}$  is a strong  $M$ -tensor and  $b \in \mathbb{R}_+^n$ . Then the sequence  $\{y_k\}$  generated by Algorithm 2 is bounded.*

*Proof.* First, similar to the proof of Lemma 4.3, it is not difficult to show that the sequence  $\{t_k y_k\}$  is bounded.

Case (i),  $\{t_k y_k\} \rightarrow 0$ . Since  $\{\theta(y_k, t_k)\} \rightarrow 0$ , we immediately have  $\{E(y_k)\} \rightarrow 0$ . Denote  $\mu_k = \|y_k\|_\infty$ ,  $\tilde{y}_k = \mu_k^{-1} y_k$  and  $\tilde{b}_k = \mu_k^{-1} b$ . Clearly, the sequence  $\{\tilde{y}_k\}$  is bounded. If  $\{y_k\}$  is unbounded, then there is a subsequence  $\{\mu_k\}_K \rightarrow \infty$ , and hence  $\{\tilde{b}_k\}_K \rightarrow 0$ . Without loss of generality, we suppose that the subsequence  $\{\tilde{y}_k\}_K$  converges to some  $\tilde{y} \geq 0$ . Denote by  $J$  the set of indices  $i$  satisfying  $\tilde{y}_i > 0$ . Obviously,  $J \neq \emptyset$ .

For some  $i \in J$ , satisfies  $y_i = \|y_k\|_\infty$ , we have

$$\begin{aligned} |\bar{E}_i(y_k, t_k)| &= \left| \frac{1}{(y_k)_i} \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \left( (y_k)_{i_2}^{\frac{1}{m-1}} \dots (y_k)_{i_m}^{\frac{1}{m-1}} \right) - \frac{b_i}{(y_k)_i} + t_k (y_k)_i \right| \\ &= \left| \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \left( (\tilde{y}_k)_{i_2}^{\frac{1}{m-1}} \dots (\tilde{y}_k)_{i_m}^{\frac{1}{m-1}} \right) - (\tilde{b}_k)_i + t_k (y_k)_i \right|. \end{aligned}$$

Taking limits in both sizes of the equality as  $k \rightarrow \infty$  with  $k \in K$  yields

$$0 = \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \left( \tilde{y}_{i_2}^{\frac{1}{m-1}} \dots \tilde{y}_{i_m}^{\frac{1}{m-1}} \right) = \sum_{i_2, \dots, i_m \in J} a_{ii_2 \dots i_m} \left( \tilde{y}_{i_2}^{\frac{1}{m-1}} \dots \tilde{y}_{i_m}^{\frac{1}{m-1}} \right),$$

Let  $\mathcal{A}_J$  be the principal subtensor of  $\mathcal{A}$  with elements  $a_{i_1 i_2 \dots i_m}$ ,  $\forall i_1, i_2, \dots, i_m \in J$ . It is a strong  $M$ -tensor but  $\mathcal{A}_J \left( \tilde{y}^{\lfloor \frac{1}{m-1} \rfloor} \right)_J^{m-1} = 0$  with  $\tilde{y} \neq 0$ . It is a contradiction. Consequently,  $\{y_k\}$  is bounded.

Case (ii), there is at least one  $i$  such that  $\liminf_{k \rightarrow \infty} t_k (y_k)_i > 0$ . In other words, there is a subsequence  $\{t_k y_k\}_K \rightarrow \tilde{y} \geq 0$  such that  $\tilde{y}_i > 0$  for at least one  $i$ . Again, denote by  $J$  the set of indices for satisfying  $\tilde{y}_i > 0$ . Since  $\{t_k\} \rightarrow 0$ , it is easy to see that

$$\lim_{k \rightarrow \infty, k \in K} (y_k)_i = +\infty, \quad \forall i \in J.$$

Denote  $\tilde{y}_k = t_k y_k$ . Similar to Case (i), we can get We derive for any  $i \in J$

$$\begin{aligned} |\bar{E}_i(y_k, t_k)| &= \left| \frac{1}{(y_k)_i} \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \left( (y_k)_{i_2}^{\frac{1}{m-1}} \dots (y_k)_{i_m}^{\frac{1}{m-1}} \right) - \frac{b_i}{(y_k)_i} + t_k (y_k)_i \right| \\ &= \left| \frac{1}{(\tilde{y}_k)_i} \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \left( (\tilde{y}_k)_{i_2}^{\frac{1}{m-1}} \dots (\tilde{y}_k)_{i_m}^{\frac{1}{m-1}} \right) - \frac{b_i}{(y_k)_i} + (\tilde{y}_k)_i \right|. \end{aligned}$$

Taking limits in both sizes of the equality as  $k \rightarrow \infty$  with  $k \in K$  yields

$$0 = \sum_{i_2, \dots, i_m} a_{ii_2 \dots i_m} \left( \tilde{y}_{i_2}^{\frac{1}{m-1}} \dots \tilde{y}_{i_m}^{\frac{1}{m-1}} \right) + \tilde{y}_i = \sum_{i_2, \dots, i_m \in J} a_{ii_2 \dots i_m} \left( \tilde{y}_{i_2}^{\frac{1}{m-1}} \dots \tilde{y}_{i_m}^{\frac{1}{m-1}} \right) + \tilde{y}_i, \forall i \in J.$$

It contradicts Theorem 2.3 (ii). □

Similar to Theorem 3.3 of [13], we have the following theorem.

**Theorem 4.6.** *Let the conditions in Assumption 2 hold, then the sequence of iterates  $\{t_k, y_k\}$  generated by Algorithm 2 converges to a positive solution of the equation 4.1 and the convergence rate is quadratic.*

## 5 Numerical Results

In this section, we do numerical experiments to test the effectiveness of the proposed methods. We implemented our methods in Matlab R2015b and ran the codes on a personal computer with 2.30 GHz CPU and 8.0 GB RAM. We used a tensor toolbox [1] to perform tensor computations.

While doing numerical experiments, similar to [12, 13], we solved the tensor equation

$$\hat{F}(x) = \hat{A}x^{m-1} - \hat{b} = 0$$

instead of the tensor equation (1.1), where  $\hat{A} := \mathcal{A}/\omega$  and  $\hat{b} := b/\omega$  with  $\omega$  being the largest value among the absolute values of components of  $\mathcal{A}$  and  $b$ . The stopping criterion is set to

$$\|\hat{F}(x_k)\| \leq 10^{-10}.$$

or the number of iterations reaches 300. The latter case means that the method is failure for the problem.

**Problem 1.** [8] We solve tensor equation (1.1) where  $\mathcal{A}$  is a symmetric strong M-tensor of order  $m$  ( $m = 3, 4, 5$ ) in the form  $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ , where tensor  $\mathcal{B}$  is symmetric whose entries are uniformly distributed in  $(0, 1)$ , and

$$s = (1 + 0.01) \cdot \max_{i=1,2,\dots,n} (\mathcal{B}\mathbf{e}^{m-1})_i,$$

where  $\mathbf{e} = (1, 1, \dots, 1)^T$ .

**Problem 2.** [32] We solve tensor equation (1.1) where  $\mathcal{A}$  is a symmetric strong M-tensor of order  $m$  ( $m = 3, 4, 5$ ) in the form  $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ , and tensor  $\mathcal{B}$  is a nonnegative tensor with

$$b_{i_1 i_2 \dots i_m} = |\sin(i_1 + i_2 + \dots + i_m)|,$$

and  $s = n^{m-1}$ .

**Problem 3.** [8] Consider the ordinary differential equation

$$\frac{d^2 x(t)}{dt^2} = -\frac{GM}{x(t)^2}, \quad t \in (0, 1),$$

with Dirichlet's boundary conditions

$$x(0) = c_0, \quad x(1) = c_1,$$

where  $G \approx 6.67 \times 10^{-11} Nm^2/kg^2$  is the gravitational constant and  $M \approx 5.98 \times 10^{24} kg$  is the mass of the earth.

Discretizing the above equation, we obtain

$$\begin{cases} x_1^3 = c_0^3, \\ 2x_i^3 - x_i^2 x_{i-1} - x_i^2 x_{i+1} = \frac{GM}{(n-1)^2}, \quad i = 2, 3, \dots, n-1, \\ x_n^3 = c_1^3. \end{cases}$$

It is a tensor equation, i.e.,

$$\mathcal{A}x^3 = b,$$

where  $\mathcal{A}$  is a 4-th order M tensor whose entries are

$$\begin{cases} a_{1111} = a_{nnnn} = 1, \\ a_{iii} = 2, \quad i = 2, 3, \dots, n-1, \\ a_{i(i-1)ii} = a_{ii(i-1)i} = a_{iii(i-1)} = -1/3, \quad i = 2, 3, \dots, n-1, \\ a_{i(i+1)ii} = a_{ii(i+1)i} = a_{iii(i+1)} = -1/3, \quad i = 2, 3, \dots, n-1, \end{cases}$$

and  $b$  is a positive vector with

$$\begin{cases} b_1 = c_0^3, \\ b_i = \frac{GM}{(n-1)^2}, \quad i = 2, 3, \dots, n-1, \\ b_n = c_1^3. \end{cases}$$

**Problem 4.** [18] We solve tensor equation (1.1) where  $\mathcal{A}$  is a non-symmetric strong M-tensor of order  $m$  ( $m = 3, 4, 5$ ) in the form  $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ , and tensor  $\mathcal{B}$  is a nonnegative tensor whose entries are uniformly distributed in  $(0, 1)$ . The parameter  $s$  is set to

$$s = (1 + 0.01) \cdot \max_{i=1,2,\dots,n} (\mathcal{B}\mathbf{e}^{m-1})_i.$$

**Problem 5.** We solve tensor equation (1.1) where  $\mathcal{A}$  is a lower triangular strong M-tensor of order  $m$  ( $m = 3, 4, 5$ ) in the form  $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ , and tensor  $\mathcal{B}$  is a strictly lower triangular nonnegative tensor whose entries are uniformly distributed in  $(0, 1)$ . The parameter  $s$  is set to

$$s = (1 - 0.5) \cdot \max_{i=1,2,\dots,n} (\mathcal{B}\mathbf{e}^{m-1})_i.$$

Since  $\mathcal{B}$  is a strictly lower triangle triangular nonnegative tensor, it has zero eigenvalue only. As a result, the spectral radius  $\rho(\mathcal{A}) = sI - \mathcal{B} = s > 0$ . It means that  $\mathcal{A}$  in Problem 5 is a strong M-tensor.

For Problem 4 and 5, we need to semi-symmetrize the tensor  $\mathcal{A}$ , i.e., find a semi-symmetric tensor  $\tilde{\mathcal{A}}$  such that

$$\mathcal{A}x^{m-1} = \tilde{\mathcal{A}}x^{m-1}.$$

The time of semi-symmetrizing the tensor is not included in CPU time.

We first test the performance of the Inexact Newton method. We set the start point  $x_0 = \varepsilon\mathbf{e}$ , where parameter  $\varepsilon$  is selected to satisfy  $f(y_0) < b$ . We set the parameter  $\sigma = 0.1$  and  $\rho = 0.5$ . And  $b$  is uniformly distributed in  $(0, 1)$  except for  $b$  in the problem 3.

For the stability of numerical results, we test the problems of different sizes. For each pair  $(m, n)$ , we randomly generate 100 tensors  $\mathcal{A}$  and  $b$ . In order to test the effectiveness of the proposed method, we compare Inexact Newton method with the QCA method in [13]. We take parameters  $\delta = 0.5, \gamma = 0.8, \sigma = 0.2, \bar{t} = 2/(5\gamma)$  as the same as in [13]. The results are listed in Tables 1, where

$$\text{IR} = \frac{\text{the number of iteration steps of the Inexact Newton method}}{\text{the number of iteration steps of the QCA method}}$$

and

$$\text{TR} = \frac{\text{the CPU time used by the Inexact Newton method}}{\text{the CPU time used by the QCA method}}.$$

We then test the effectiveness of the Regularized Newton method. We set the initial point  $x_0 = 0.1 * \mathbf{e}$  and  $b \in \mathbb{R}_+^n$  has zero elements except the problem 3. We first generate a vector  $b^0 \in \mathbb{R}^n$  whose elements are uniformly distributed in  $(0, 1)$ , then we set

$$b_i = \begin{cases} b_i^0, & \text{if } b_i^0 \leq 0.6, \\ 0, & \text{if } b_i^0 > 0.6. \end{cases}$$

to get a vector  $b \in \mathbb{R}_+^n$ . In order to get the positive solution of the problem 5, the first component of vector  $b$  can't be equal to 0, so we set the first component  $b_1 = 0.1$ .

Table 1: Comparison between Inexact Newton method and QCA method with  $b \in \mathbb{R}_{++}^n$ .

	$(m, n)$	(3,10)	(3,100)	(3,300)	(3,500)	(4,10)	(4,50)	(4,100)	(5,10)	(5,30)
IR	Problem 1	89.2%	91.5%	91.5%	91.0%	93.0%	93.7%	94.3%	95.2%	96.3%
	Problem 2	91.0%	91.4%	90.2%	90.5%	95.7%	94.8%	93.1%	97.2%	96.2%
	Problem 3	-	-	-	-	11.1%	9.1%	8.3%	-	-
	Problem 4	91.8%	91.2%	91.3%	90.5%	94.4%	94.7%	93.2%	97.1%	93.9%
	Problem 5	89.8%	90.4%	89.6%	89.3%	95.2%	91.5%	93.0%	95.1%	95.6%
TR	Problem 1	48.0%	66.6%	87.7%	88.2%	67.3%	92.3%	94.0%	80.0%	97.0%
	Problem 2	50.0%	73.8%	88.8%	88.8%	67.4%	94.0%	93.7%	79.0%	96.6%
	Problem 3	-	-	-	-	20.3%	15.4%	14.2%	-	-
	Problem 4	54.1%	73.3%	89.6%	89.6%	66.0%	94.7%	94.1%	74.6%	95.4%
	Problem 5	45.7%	74.1%	87.4%	88.5%	59.1%	92.3%	95.0%	74.6%	99.4%

We compare the Regularized Newton Method with QCA method. We take the parameters  $\sigma = 0.1, \rho = 0.8, \gamma = 0.9$  and  $\bar{t} = 0.01$  in Regularized Newton Method and the parameters in QCA method are the same as above. The results are listed in Table 2, where

$$\text{IR} = \frac{\text{the number of iteration steps of the Regularized Newton method}}{\text{the number of iteration steps of the QCA method}}$$

and

$$\text{TR} = \frac{\text{the CPU time used by the Regularized Newton method}}{\text{the CPU time used by the QCA method}}.$$

Table 2: Comparison between Regularized Newton method and QCA method with  $b \in \mathbb{R}_{++}^n$ .

	$(m, n)$	(3,10)	(3,100)	(3,300)	(3,500)	(4,10)	(4,50)	(4,100)	(5,10)	(5,30)
IR	Problem 1	92.4%	59.7%	71.6%	67.4%	93.2%	67.2%	59.5%	95.7%	78.4%
	Problem 2	83.9%	61.5%	50.3%	49.4%	89.1%	56.0%	59.3%	87.0%	59.4%
	Problem 3	-	-	-	-	83.3%	80.0%	81.0%	-	-
	Problem 4	94.3%	65.8%	58.1%	60.9%	95.1%	67.7%	53.8%	95.5%	76.5%
	Problem 5	80.0%	81.0%	81.1%	81.4%	75.4%	77.6%	76.4%	72.4%	74.0%
TR	Problem 1	80.0%	78.7%	89.2%	82.4%	93.2%	77.2%	71.6%	86.8%	97.8%
	Problem 2	72.7%	72.9%	61.3%	60.4%	87.5%	61.4%	64.0%	94.1%	65.3%
	Problem 3	-	-	-	-	76.5%	97.6%	98.1%	-	-
	Problem 4	80.6%	81.3%	65.7%	76.7%	92.3%	78.1%	65.0%	97.4%	99.6%
	Problem 5	72.1%	94.9%	89.6%	91.0%	86.9%	79.3%	77.7%	82.5%	80.8%

The data in Table 1 and 2 show that for all test problems the Inexact Newton method and the Regularized Newton method are better than QCA method in terms of the number of iterations and CPU time. It is worth noting that although the QCA method in [13] does not established the convergence property in the case of  $b \in \mathbb{R}_+^n$ , we find that in the case of  $b \in \mathbb{R}_{++}^n$ , the QCA method can still find the solution of the problem successfully. For the convenience of readers, we only list the relative results. More detailed numerical results can be found in the Appendix.

## 6 Conclusions

In this paper, we have proposed an inexact Newton method for solving M-tensor equations with positive constant term. We have shown that if the constant term of the equation is nonnegative, then we can directly determine the zero components of the nonnegative solution of M-tensor equation. As a result, finding a nonnegative solution of an M-tensor equation is transformed into finding a positive solution of a lower dimensional M-tensor equation. Based on this property, we have extended this method to the M-tensor equations with nonnegative constant term. Both methods possess global and quadratic convergence. Numerical results indicate that the proposed methods are effective and competitive.

## References

- [1] B.W. Bader, T.G. Kolda and others, MATLAB Tensor Toolbox Version 2.6 (2015)
- [2] X. Bai, H. He, C. Ling C and G. Zhou, A nonnegativity preserving algorithm for multilinear systems with nonsingular M-tensors, *Numer. Algorithms* 87 (2021) 1301–1320
- [3] A. Berman and R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, SIAM, Philadelphia, 1994
- [4] H. Bozorgmanesh, M. Hajarain and A.T. Chronopoulos, Interval Tensors and their application in solving multi-linear systems of equations, *Comput. Math. Appl.* 79(3) (2020) 697-715
- [5] M. Brazell, N. Li, C. Navasca and C. Tamon, Solving multilinear systems via tensor inversion, *SIAM J. Matrix Anal. Appl.* 34 (2013) 542-570
- [6] K.C. Chang, K. Pearson and T. Zhang, Primitivity, the convergence of the NQZ method, and the largest eigenvalue for nonnegative tensors, *SIAM. J. Matrix Anal. Appl.* 32 (2011) 806-819
- [7] W. Ding, L. Qi L and Y. Wei, M-tensors and nonsingular M-tensors, *Linear Algebra Appl.* 10 (2013) 3264-3278
- [8] W. Ding, Y. Wei, Solving multi-linear systems with M-tensors, *J. Sci. Comput.* 68 (2016) 689-715
- [9] F. Facchinei and C. Kanzow, Beyond monotonicity in regularization methods for non-linear complementarity problems, *SIAM J. Control Optim.* 37 (1999) 1150-1161
- [10] H.Y. Fan, L. Zhang, E.K. Chu and Y. Wei, Numerical solution to a linear equation with tensor product structure, *Numeri. Linear Algebra Appl.* 24 (2017) 1-22
- [11] M.S. Gowda, Z. Luo, L. Qi and N. Xiu, Z-tensor and complementarity problems, (2015) arXiv:1510.07933
- [12] L. Han, A homotopy method for solving multilinear systems with M-tensors, *Appl. Math. Lett.* 69 (2017) 49-54
- [13] H. He, C. Ling, L. Q and G. Zhou, A globally and quadratically convergent algorithm for solving multilinear systems with M-tensors, *J. Sci. Comput.* 76 (2018) 1718-1741

- [14] D. Kressner D and C. Tobler, Krylov subspace methods for linear systems with tensor product structure, *SIAM J. Matrix Anal. Appl.* 31 (2010) 1688-1714
- [15] D.H. Li, S. Xie and H.R. Xu, Splitting methods for tensor equations, *Numeri. Linear Algebra Appl.* (2017) <https://doi.org/10.1002/nla.2102>
- [16] X. Li and M.K. Ng, Solving sparse non-negative tensor equations: algorithms and applications, *Front. Math. China.* 10 (2015) 649-680
- [17] Z. Li, Y.-H. Dai and H. Gao, Alternating projection method for a class of tensor equations, *J. Comput. Appl. Math.* 346 (2019) 490-504
- [18] D.H. Li, H.B. Guan and X.Z. Wang, Finding a nonnegative solution to an M-tensor equation, *Pac. J. Optim.* 16(3) (2020) 419-440
- [19] L.H. Lim, Singular values and eigenvalues of tensors, a variational approach, *In: Proceedings of the 1st IEEE International Workshop on Computational Advances of Multi-tensor Adaptive Processing*, 1 (2005) 129-132
- [20] D. Liu, W. Li and S.W. Vong, The tensor splitting with application to solve multi-linear systems, *J. Comput. Appl. Math.* 330 (2018) 75-94
- [21] D. Liu, W. Li, S.W. Vong, A new preconditioned SOR method for solving multilinear systems with an M-tensor, *Calcolo* (2020) <http://doi.org/10.1007/s10092-020-00364-8>
- [22] W. Li, D. Liu and S.W. Vong, Comparison results for splitting iterations for solving multi-linear systems, *Appl. Numer. Math.* 134 (2018) 105-121
- [23] M. Liang, B. Zheng and R. Zhao, Alternating iterative methods for solving tensor equations with applications, *Numer. Algorithms.* 80(4) (2019) 1437-1465
- [24] C.Q. Lv and C.F. Ma, A Levenberg-Marquardt method for solving semi-symmetric tensor equations, *J. Comput. Appl. Math.* 332 (2018) 13-25
- [25] L. Qi, Eigenvalues of a real supersymmetric tensor, *J. Symbolic Comput.* 40 (2005) 1302-1324
- [26] L. Qi, H. Chen and Y. Chen, *Tensor Eigenvalues and Their Applications*, Springer, 2018
- [27] L. Qi and Z. Luo, *Tensor Analysis, Spectral Theory and Special Tensors*, SIAM Philadelphia, 2017
- [28] X. Wang, M. Che and Y. Wei Y, Neural networks based approach solving multi-linear systems with M-tensors, *Neurocomputing*, 351 (2019) 33-42
- [29] X. Wang, M. Che and Y. Wei, Neural network approach for solving nonsingular multi-linear tensor systems, *J. Comput. Appl. Math.* (2020) <http://doi.org/10.1016/j.cam.2019.112569>
- [30] X. Wang, M. Che M and Y. Wei, Existence and uniqueness of positive solution for  $H^+$ -tensor equations, *Appl. Math. Lett.* 98 (2019) 191-198
- [31] Z.J. Xie, X.Q. Jin and Y.M. Wei, A fast algorithm for solving circulant tensor systems, *Linear Multilinear Algebra*, 65 (2017) 1894-1904 (2017)

- [32] Z.J. Xie, X.Q. Jin and Y.M. Wei, Tensor methods for solving symmetric M-tensor systems, *J. Sci. Comput.* 74 (2018) 412-425
- [33] Y. Xu, W. Gu and Z.H. Huang, Properties of the nonnegative solution set of multi-linear equations, *Pac. J. Optim.* 15 (2019) 441-456
- [34] W. Yan, C. Ling, L. Ling and H. He, Generalized tensor equations with leading structured tensors, *Appl. Math. Comput.* 361 (2019) 311-324
- [35] Y. Zhang Y, Q. Liu and Z. Chen, Preconditioned Jacobi type method for solving multi-linear systems with M-tensors, *Appl. Math. Lett.* (2020). <http://doi.org/10.1016/j.aml.2020.106287>
- [36] L. Zhang, L. Qi and G. Zhou G, M-tensor and some applications, *SIAM J. Matrix Anal. Appl.* 35 (2014) 437-452

## Appendix

In this section, we list the detailed numerical results of the proposed methods compared with QCA method. The results are listed in Tables 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12, where the columns 'Iter', 'Time', 'Res' and 'Ls-iter' stand for the total number of iterations, the computational time (in second) used for the method, the residual  $\|\hat{\mathcal{A}}x_k^{(m-1)} - \hat{b}\|$  and the total number of iterations of linear search.

Table 3: Comparison between Inexact Newton method and QCA method on Problem 1 .

$(m, n)$	Inexact Newton method				QCA			
	Iter	Time	Res	Ls-iter	Iter	Time	Res	Ls-iter
(3,10)	6.6	0.00024	8.5E-12	0	7.4	0.00050	6.1E-12	0
(3,100)	9.7	0.00829	9.0E-12	0	10.6	0.01244	1.0E-11	0
(3,300)	11.9	0.32238	9.8E-12	0	13.0	0.36766	1.7E-11	0
(3,500)	12.1	1.44961	5.1E-12	0	13.3	1.64275	7.4E-12	0
(4,10)	6.6	0.00033	5.0E-12	0	7.1	0.00049	9.2E-12	0
(4,50)	8.9	0.04552	1.2E-11	0	9.5	0.04931	1.4E-11	0
(4,100)	10.0	0.77240	1.3E-11	0	10.6	0.82138	6.7E-12	0
(5,10)	6.0	0.00052	1.3E-11	0	6.3	0.00065	1.3E-11	0
(5,30)	7.9	0.15599	8.9E-12	0	8.2	0.16078	1.2E-11	0

Table 4: Comparison between Inexact Newton method and QCA method on Problem 2.

$(m, n)$	Inexact Newton method				QCA			
	Iter	Time	Res	Ls-iter	Iter	Time	Res	Ls-iter
(3,10)	7.1	0.00020	1.0E-11	0	7.8	0.00040	4.7E-12	0
(3,100)	9.6	0.00899	5.7E-12	0	10.5	0.01218	9.1E-12	0
(3,300)	11.9	0.32718	8.6E-12	0	13.2	0.36855	1.1E-11	0
(3,500)	12.4	1.49714	8.0E-12	0	13.7	1.68625	8.3E-12	0
(4,10)	6.7	0.00029	8.6E-12	0	7.0	0.00043	9.8E-12	0
(4,50)	9.1	0.04654	7.8E-12	0	9.6	0.04950	1.4E-11	0
(4,100)	9.5	0.74050	1.5E-11	0	10.2	0.79047	1.7E-11	0
(5,10)	6.9	0.00049	6.5E-12	0	7.1	0.00062	1.3E-11	0
(5,30)	7.6	0.15073	1.0E-11	0	7.9	0.15600	1.4E-11	0

Table 5: Comparison between Inexact Newton method and QCA method on Problem 3.

$(m, n)$	Inexact Newton method				QCA			
	Iter	Time	Res	Ls-iter	Iter	Time	Res	Ls-iter
(4,10)	1.0	0.00012	9.2E-15	1.0	9.0	0.00059	6.4E-12	1.0
(4,50)	1.0	0.00947	2.0E-15	1.0	11.0	0.06154	4.2E-14	1.0
(4,100)	1.0	0.14564	2.1E-15	1.0	12.0	1.02232	1.9E-15	1.0

Table 6: Comparison between Inexact Newton method and QCA method on Problem 4.

$(m, n)$	Inexact Newton method				QCA			
	Iter	Time	Res	Ls-iter	Iter	Time	Res	Ls-iter
(3,10)	6.7	0.00020	8.8E-12	0	7.3	0.00037	9.6E-12	0
(3,100)	10.3	0.00934	7.9E-12	0	11.3	0.01274	1.1E-11	0
(3,300)	11.6	0.31909	1.2E-11	0	12.7	0.35600	1.3E-11	0
(3,500)	12.4	1.50356	7.9E-12	0	13.7	1.67812	8.5E-12	0
(4,10)	6.8	0.00031	3.7E-12	0	7.2	0.00047	8.9E-12	0
(4,50)	8.9	0.04571	1.4E-11	0	9.4	0.04826	1.1E-11	0
(4,100)	9.6	0.74759	1.4E-11	0	10.3	0.79482	1.4E-11	0
(5,10)	6.6	0.00047	5.4E-12	0	6.8	0.00063	1.0E-11	0
(5,30)	7.7	0.15334	1.5E-11	0	8.2	0.16067	1.5E-11	0

Table 7: Comparison between Inexact Newton method and QCA method on Problem 5.

$(m, n)$	Inexact Newton method				QCA			
	Iter	Time	Res	LS-iter	Iter	Time	Res	LS-iter
(3,10)	7.9	0.00016	7.2E-12	0.4	8.8	0.00035	4.7E-12	0.6
(3,100)	10.3	0.00758	1.2E-11	0.1	11.4	0.01023	9.5E-12	0.6
(3,300)	12.1	0.27135	1.2E-11	0	13.5	0.31054	8.8E-12	0.5
(3,500)	12.5	1.44273	1.4E-11	0	14.0	1.63077	1.7E-11	0.4
(4,10)	8.0	0.00026	4.2E-12	0.5	8.4	0.00044	1.1E-11	0.7
(4,50)	9.7	0.04921	8.8E-12	0.2	10.6	0.05329	8.2E-12	0.5
(4,100)	10.6	0.82689	7.0E-12	0.2	11.4	0.87036	1.5E-11	0.7
(5,10)	7.7	0.00047	6.3E-12	0.5	8.1	0.00063	1.0E-11	0.7
(5,30)	8.6	0.17388	6.0E-12	0.4	9.0	0.17493	1.4E-11	0.6

Table 8: Comparison between Regularized Newton method and QCA method on Problem 1.

$(m, n)$	Regularized Newton method				QCA			
	Iter	Time	Res	LS-iter	Iter	Time	Res	LS-iter
(3,10)	7.3	0.00036	4.3E-12	0.0	7.9	0.00045	7.1E-12	0.0
(3,100)	4.6	0.00734	1.2E-11	0.6	7.7	0.00933	7.5E-12	4.8
(3,300)	5.3	0.19312	9.8E-12	0.8	7.4	0.21646	1.1E-11	15.2
(3,500)	6.0	0.97474	1.4E-11	0.9	8.9	1.18350	2.1E-11	18.0
(4,10)	8.2	0.00055	7.4E-12	0.0	8.8	0.00059	7.9E-12	0.0
(4,50)	4.3	0.02707	8.4E-12	0.6	6.4	0.03507	1.1E-11	3.6
(4,100)	5.0	0.47484	1.3E-11	0.8	8.4	0.66361	3.1E-11	14.3
(5,10)	8.9	0.00079	8.3E-12	0.0	9.3	0.00091	8.8E-12	0.0
(5,30)	4.0	0.11252	1.5E-12	1.0	5.1	0.11508	1.8E-11	1.5

Table 9: Comparison between Regularized Newton method and QCA method on Problem 2.

$(m, n)$	Regularized Newton method				QCA			
	Iter	Time	Res	LS-iter	Iter	Time	Res	LS-iter
(3,10)	5.2	0.00024	6.7E-12	0.8	6.2	0.00033	6.8E-12	3.6
(3,100)	6.7	0.00894	7.9E-12	0.8	10.9	0.01227	1.1E-11	35.4
(3,300)	7.2	0.25393	1.0E-11	1.0	14.3	0.41420	2.2E-12	69.1
(3,500)	7.9	1.22181	1.0E-11	0.8	16.0	2.02200	1.0E-12	87.0
(4,10)	4.9	0.00035	1.2E-11	0.7	5.5	0.00040	1.6E-11	3.8
(4,50)	6.5	0.03804	1.2E-11	0.8	11.6	0.06193	8.4E-13	42.3
(4,100)	7.0	0.61673	1.2E-11	0.8	11.8	0.96383	1.3E-12	50.2
(5,10)	4.7	0.00048	1.3E-11	0.7	5.4	0.00051	7.0E-12	3.5
(5,30)	6.0	0.13580	1.3E-11	0.7	10.1	0.20807	1.7E-12	31.6

Table 10: Comparison between Regularized Newton method and QCA method on Problem 3.

$(m, n)$	Regularized Newton method				QCA			
	Iter	Time	Res	Ls-iter	Iter	Time	Res	Ls-iter
(4,10)	15.0	0.00101	6.9E-16	0	18.0	0.00132	1.6E-12	0
(4,50)	16.0	0.09657	3.7E-11	0	20.0	0.09892	3.9E-12	0
(4,100)	17.0	1.58356	9.0E-12	0	21.0	1.61399	8.8E-14	0

Table 11: Comparison between Regularized Newton method and QCA method on Problem 4.

$(m, n)$	Regularized Newton method				QCA			
	Iter	Time	Res	Ls-iter	Iter	Time	Res	Ls-iter
(3,10)	6.6	0.00029	7.4E-12	0.0	7.0	0.00036	6.7E-12	0.0
(3,100)	4.8	0.00669	1.4E-11	0.6	7.3	0.00823	4.6E-12	5.8
(3,300)	5.4	0.18935	1.2E-11	0.5	9.3	0.28820	5.9E-11	17.9
(3,500)	5.6	0.91665	9.6E-12	0.8	9.2	1.19580	2.0E-11	20.0
(4,10)	7.8	0.00048	8.1E-12	0.0	8.2	0.00052	6.9E-12	0.0
(4,50)	4.4	0.02734	9.8E-12	0.6	6.5	0.03502	1.0E-11	4.0
(4,100)	5.0	0.46779	1.4E-11	0.7	9.3	0.72021	3.7E-11	17.0
(5,10)	8.5	0.00075	8.8E-12	0.0	8.9	0.00077	1.0E-11	0.0
(5,30)	3.9	0.10700	1.6E-11	1.2	5.1	0.10740	1.2E-11	1.7

Table 12: Comparison between Regularized Newton method and QCA method on Problem 5.

$(m, n)$	Regularized Newton method				QCA			
	Iter	Time	Res	Ls-iter	Iter	Time	Res	Ls-iter
(3,10)	8.0	0.00031	2.0E-12	0.5	10.0	0.00043	6.9E-12	12.8
(3,100)	11.9	0.01314	5.5E-12	0.8	14.7	0.01385	5.3E-11	73.4
(3,300)	14.2	0.36332	2.1E-11	0.7	17.5	0.40539	1.0E-11	109.4
(3,500)	15.3	2.02149	2.0E-11	0.5	18.8	2.22143	4.6E-12	127.9
(4,10)	8.6	0.00053	4.1E-12	0.5	11.4	0.00061	1.3E-11	26.7
(4,50)	12.5	0.06472	1.6E-11	0.8	16.1	0.08161	3.1E-12	95.1
(4,100)	14.9	1.16723	1.6E-11	0.8	19.5	1.50176	5.2E-13	144.7
(5,10)	9.2	0.00080	6.9E-12	0.5	12.7	0.00097	6.3E-12	44.3
(5,30)	12.8	0.27238	1.4E-11	1.8	17.3	0.33700	3.9E-13	115.6

From the data in the Tables 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12, we can see that the proposed methods are effective for all test problems. In terms of the number of iterations and CPU time, Inexact Newton method and Regularized Newton method are better than QCA method, and the number of linear search of the Regularized Newton method are far less than that of the QCA method.

*Manuscript received 22 June 2021*  
*revised 30 August 2021*  
*accepted for publication 30 September 2021*

DONG-HUI LI

School of Mathematical Sciences, South China Normal University  
Guangzhou, 510631, China  
E-mail address: lidonghui@m.scnu.edu.cn

HONG-BO GUAN

School of Mathematical Sciences, South China Normal University  
Guangzhou, 510631, China;  
School of Mathematical Sciences and Energy Engineering  
Hunan Institute of Technology, HengYang, 421002, China  
E-mail address: hongbo.guan@m.scnu.edu.cn

JIE-FENG XU

School of Mathematical Sciences, South China Normal University  
Guangzhou, 510631, China  
E-mail address: 2018021699@m.scnu.edu.cn